DOCUMENT RESUME

ED 209 764 EA 014 227

AUTHOR Foley, Walter J.
TITLE On Evaluation and the Evaluation of Teachers.
INSTITUTION Iowa Univ., Iowa City. Inst. for School Executives.
PUB DATE Apr 81
NOTE 5p.
AVAILABLE FROM Institute for School Executives, University of Iowa, 210 Lindquist Center, Iowa City, IA 52242 ($1.00).
JOURNAL CIT Executive Review; v1 n5 Apr 1981

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Administrator Role; Elementary Secondary Education; Evaluation Criteria; *Evaluation Methods; Evaluation Needs; *Teacher Evaluation; *Teacher Improvement; Teacher Supervision

ABSTRACT

Seeking to clarify the purpose of teacher evaluation, the author advances a strategy in which improvement of instruction is the primary goal of evaluation. The paper divides the activities and strategies involved in evaluation into three broad categories: observation, judgment, and assignment of value. It recommends a process in which only those teachers most in need of improvement are identified and evaluated. This evaluation is followed by implementation of specific measures that focus on improving individual behaviors. The author also discusses the role of consensus and agreement over the purpose and nature of the evaluation process, the reliability and validity of the measurement standards employed, and student learning. The author concludes that the primary task in evaluating teachers is to finish with statements that relate to an acceptable theory of instruction rather than to teachers' personal characteristics. (JEH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Volume 1, Number 5
April 1981

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

W. Bozeman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Executive Review

Approved KCP

Institute for School Executives — The University of Iowa

ED209764

EA 014 227

ON EVALUATION AND THE EVALUATION OF TEACHERS

Walter J. Foley

Our concern is with the evaluation of teachers. This concern will be addressed by beginning with a definition, elaborating on the meaning of that definition as it relates to teaching, discussing the problems encountered when attempting to evaluate teachers, and concluding with some suggesions. The discussion focuses on the unique place of the improvement of instruction as a purpose for the evaluation of teachers. The discussion, while mentioning briefly other purposes for teacher evaluation such as dismissal and merit, does not center on these other purposes.

Part I

For most people *evaluation* has two definitions. One is based upon a value that has already been placed on whatever is of interest. Consider this statement: Mary is a good teacher. "*Good*" is an adjective that, in an evaluative way, characterizes Mary as a teacher. The other definition centers on the activities that concluded with the assignment of that characteristic to Mary. What activities are involved in order to be able to end the process by stating that Mary is "good"? What does one need to know in order to be able to say that Mary is a "good" teacher? What do we have to know in order to reach this judgment? The process of arriving at a judgment of worth constitutes the second definition of evaluation. This is to say that the problems encountered in teacher evaluation center on the processes that occur before we are able to state with confidence our judgment of Mary's value or worth as a teacher.

First, assume that the process is over and that an evaluative statement has been made (value was assigned). We tend to see the value included as an ascribed characteristic. Ascribed in the sense that the value is a property that characterizes Mary as a teacher. When viewed as a value statement, this example presents our evaluation as a property that is characteristic of Mary. Our intended meaning falls under the first definition. Our evaluation of Mary as a good teacher is accepted by others simply because we have said that it represents the fact of the matter. The procedural questions contained in the second definition come into play when we do not accept what has been stated. For example, how do we know that what we said is to be believed? Our concern with evaluation as an activity centers on how well we are able to establish that our evaluative statements are factual. In this paper, it will be argued that the acceptance of a teacher evaluation program centers on our ability to provide an evaluation methodology that results in factual and publicly acceptable evaluative statements.

Frames of Reference

The activities and strategies involved in the evaluation process can be easily grouped within three major categories. First, we observe and compare what is relevant in what we observe to an ideal or standard, make a judgment, and make a statement that assigns a value. The outcome is the assignment of a value position to what we are interested in judging. Positions are on a scale that can range from "poorer than" through "equal to" in relation to our ideal or standard. It is interesting to note that when what is of interest exceeds an ideal, it replaces the ideal and becomes a "new" ideal. Judgments cannot be better than our idea of best. Standards, on the other hand, define only a "level of expectation." For example, a behavioral objective serves as a standard.

Second, from our experience we can compare what we observe with what we know about similar things of interest. This choice forces us to form a distribution that includes what is of interest. The observed is compared on the distribution of other known occurrences that we create. The place of the present instance on the distribution is noted and a relative value is assigned. In order to employ this method, we need to know how what is of interest is distributed.

The third category includes the use of our previous observations and experience with respect to the *same* interest in order to be able to judge a change, if any, in the present observation. The previous observations are seen as independent in the sense that they allow us to make our judgments of change or alteration by comparing the past with the present. Under this method, at least one previous experience with our interest is necessary.

To put it simply, whenever we evaluate a particular "whatever" that is of interest to us, our choice of a frame of reference is limited by our experience. We cannot refer to an ideal or standard if one is not available, a distribution cannot be constructed if unknown, and we need a "history of phenomena" in order to judge its present status.

After what is of interest is assigned (tied) to our frame of reference (given that a frame of reference is available), we can proceed with our evaluations. The activities fall under the second definition and the value that is assigned falls under the first definition of evaluation. The point is that evaluative activities end with a value being ascribed to whatever is of interest to us. For what we say we "know" (that value that we assign) to be seen as informative to others, their encounters with a similar phenomena have to provide a degree of shared expectation (a degree of likelihood-probability). When others are comfortable with and see our evaluative statements as factual, frames of reference do not have to be made explicit.

The evaluation of teachers, when reduced to its simplest elements, involves three steps: observation, comparison to a chosen frame of reference, and the assignment of a value. In one sense, the frame of reference we choose shapes what we will count. Choosing a frame of reference is also the choice of the rules that we will follow when making our judgments of worth. These rules determine what we will accept as measures of what is of interest in what is observed. It is the sharing of our judgments with others that extends the third step into the public world. When we end our evaluations and make public statements, we can expect that others will judge them as "more or less" true on the basis of how well we have followed the rules.

An Example

The outcome of any evaluative activity is dependent, in large measure, on the choice of a frame of reference that allows us to make judgments about our observations. Consider the problem Mary faces when evaluating her class:

> The "best" pupil in Mary's reading class scored 80 percent on her reading test. The ideal score on a reading test for a pupil is (to Mary) 100 percent correct. On the three previous reading tests, the best pupil scored 50 percent, 60 percent, and 70 percent correct. The classmates all scored below the best pupil's percent correct on the same sequence of tests.

How should Mary evaluate her best pupil? In relation to her ideal score, the best pupil was always low. Compared to past performance, this pupil

has improved every time and compared to the classmates, the pupil of interest has performed "best" on every occasion. Obviously, the frame of reference Mary chooses will color her evaluation of this pupil. In a sense of fairness, most people would say that the frame of reference chosen by a teacher should be made public, at least to the class.

The relationships between observations, frames of reference, and evaluative statements posited for students also holds for our evaluations of teachers. When we employ an ideal or standard, it has to be shared by at least those being evaluated. A distribution also has to be made public prior to the evaluative statements that employ this frame of reference and "change" demands that we have previous experience available. These rules seem most appropriate when the purpose of our evaluative activities is the improvement of instruction.

Evaluative Statements

Given that there are three frames of reference available to us as evaluators, choosing one does not completely solve our problems. This choice often forces us to attempt to answer a second question: "Is this frame of reference acceptable in making a 'fair' evaluation?" This question centers our attention on the public nature of the acceptance we seek when we want to share our judgments of worth with those being evaluated. Before we can be seen as fair evaluators by others, the value of our frame of reference for our purpose has to be publicly established. The rules that have been accepted for judging fairness follow the publicly accepted rules for the evenhandedness of our procedures. These are the same criteria that most people have in mind when they evaluate an evaluator. They are what most people mean when they say that a particular person is a fair and impartial judge. More will be said about criteria later in the paper.

When we engage in evaluative activities, we hope to be able to end with evaluative statements that will be seen as fair (good in an evaluative sense) by others. We hope to be able to frame simple declarative statements. This is to say that our goal as evaluators is to end our activities with clear statements that relate our activities to our purpose.

The Place of Purpose

The accuracy of our evaluative statements is tested by matching our observations and frames of reference with our purpose. Others have the problem of evaluating their factual content. Logic and mathematics provide us with insights as we attempt to understand the two-part nature of evaluative activities. When we try to establish tests for the "truth value" of our evaluative statements, we find that to support our statements of value we use our evaluative activities. The activities that lie behind our judgments are seen as more or less appropriate or internally consistent with our purpose.

To see this relationship more clearly, we often turn our evaluative questions into propositions like *if* Mary, *then* good teacher. This form allows us to go behind our evaluative statements and attempt to judge the quality of the method employed against the purpose posited. Depending on our purpose, we go about describing (defining? defending?) our judgments differently. When we are satisfied, we stop. We do not pay much attention to our purpose. In fact, we often defend the factual nature of our statements when they are questioned only by defending the method/measures employed.

Evaluative statements that are intended to improve instruction will defer from those framed in order to dismiss the incompetent or to justify decisions of merit. Experienced administrators will tell you that if only a small portion of the teaching staff significantly improved their instruction, this purpose would be served. They will also say that evaluating every teacher in order to improve the instructional practices of those who need to improve has only misdirected (wasted) a great deal of administrative time and effort. We know that in order to improve the teaching performance in a school significantly, we do not need to improve the performance of all of the teachers. Yet, when we design our evaluative procedures in ways that treat all the teachers in the same way, we end with measures that fit a rectangular distribution. This is unproductive. It is unproductive in the sense that since the time devoted to the effort ignores its purpose, we quickly reach the point of diminishing returns. The argument is that our evaluative time and effort would be spent more profitably by working with that smaller portion of the total group whose improvement (given that they could be improved) we already know would contribute significantly to the improvement of instruction. The question of "how" is addressed next. It should be kept in mind that what is presented will be considered inappropriate by those who start with the assumption that everyone needs improvement.

An Axiom

In any group, the behavior of the few disproportionately influences the behavior of the total. We already know from our experience that of the many tasks that are accomplished by teachers in a school in the course of its functioning, only a few are key to its success (Note 1). Many of our other "problems" such as tardiness, scheduling, discipline, progress reports, and so forth, can all serve as additional examples. What we know in each case is that by changing the behavior of only a few (an estimate of no more than 10 to 20%), we will disproportionately influence the total.

To serve the purpose of the improvement of instruction, we should build upon the inverse relationship between the problem of instructional improvement and the source of the problem. Consider this example: administrators follow the standard procedure of notifying each teacher of the date scheduled for observation, holding a preobservation meeting, observing (usually for no more than 30 minutes in duration), and sharing the evaluation with the teacher in a postobservation conference. The statements that are the product of this sequence and the recommendations (changes in practice, if any) are used to determine a teacher's "need for improvement." For those few that are found to be in need, an additional follow-up notification/preconference/observation/postconference sequence is often employed. One such model from my experience required 10 hours of administrator time per teacher per year.

The method costs an inordinate amount of teacher/administrator time for those teachers (the 80% or more) that are known to be satisfactory or better before the activity begins. A uniform procedure practically guarantees that not enough time will be available for the focused effort that is "critical" to achieving the purpose the program is designed to serve. We find that only 10 to 20 percent of the available time can be devoted to the teachers who pose 80 to 90 percent of the instructional problems. The net result is an evaluation methodology that is cumbersome, costly, time consuming, and irrelevant to the majority of the teachers that the method is designed to serve.

To answer the "how" question posed earlier, we have developed a methodology that better serves the evaluative purpose. The method is based upon the axiom and first focuses our efforts on the identification of that small (small in the sense of representing only a few of the many in the group) number of teachers who would benefit from an intensive effort to improve their instructional practices.

We would begin the process by ranking teachers according to some "quick and easy" method, that is, administrator/teacher/student judgment to identify those whose improvement would most likely increase the overall success of the instructional program. For example, you might list the teachers in descending order of teaching competence based on your own perceptions. Next, you would convert this list to a percent distribution with 100 percent: those most in need at the top. In terms of the midpoint of need, not the number of teachers, mark 50 percent with those with no need at the bottom of the list (Figure 1). The percent printed across the bottom on the figure represents the cumulative percent of administrative time available. The results will follow the form of the

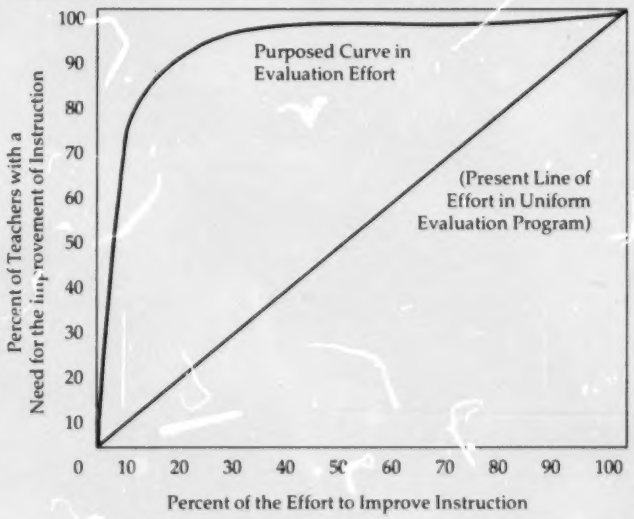

Figure 1. Cumulative Graph of the Percent of Need for Instructional Improvement and the Percent of Teachers with That Need

cumulative frequency curve. A small percentage (usually no more than 20) of the teachers would be expected to account for the vast majority (80 to 90) of the time expended on instructional improvement under the curve. Evaluative efforts would be focused on the improvement of their instructional effectiveness. For contrast, the typical uniform evaluation practice of "treat everyone the same" is shown on the diagonal of the same graph.

A Two-Step Evaluation

To the extent that you are willing to accept the assumption that the purpose of this evaluative activity fits the general shape of the cumulative frequency curve shown as Figure 1, a two-part strategy is appropriate. By also assuming that the teachers who need improvement should be helped early in the year (that is simply good practice), we would schedule the hours available for evaluative activities. The focus would be on the differences that have the best chance of making a difference in terms of the achievement of the instructional achievement purpose. The majority of your evaluative time would then be available for individual programs designed to help the teachers in the 10 to 20 percent group. Again, evaluative strategies that focus on the small portion of problem cases should have two steps. Begin with an initial screening in order to identify the members that are "key" to the achievement of the purpose. Then, provide the specific measures that are needed in order to end the activity with evaluative statements that focus on the improvement of individual behaviors. Before detailing the development of measures, we will deal with the other two common purposes posed for teacher evaluation.

The second purpose, "Evaluation for the dismissal of the incompetent," should be addressed while keeping in mind that this evaluation design strategy is best developed in consultation with a lawyer. Incompetence as the basis for termination requires a legal definition and the methods employed in the process of developing the information necessary to support this decision differ from those designed to improve instruction. A lawyer becomes the evaluation design consultant in the evaluation of competence.

To address the third purpose, "Evaluation for merit," the design would be similar to that proposed for the improvement of instruction. A two-step methodology would also apply since the axiom would be expected to hold, that is, that only a small portion of the group would be included. The two purposes focus, in a real sense, on the two ends of the teacher distribution. The strategy proposed for instructional improvement would also hold for the "merit" end of a teacher distribution for many of the same reasons that have already been discussed in support of the model proposed for instructional improvement.

Since it has been argued that the "success" of an evaluation is also, in large measure, dependent upon the public value attached to the statements made, the statements need to be seen as both effective in terms of the achievement of the purpose of the effort and efficient in terms of the use of our available evaluative resources. Up to this point, the discussion was intended to clarify the concept of evaluation as it relates to teaching, present an axiom to guide our teacher evaluation design efforts, and spell out a two-step evaluation strategy. We now turn toward other considerations.

Other Considerations

While the ideal evaluative statement presents the *invariant* conclusion of an identity between observation, frame of reference, and purpose, the reality of the case is that when we evaluate teachers, we find less than an invariant relationship. Put in other terms, when we connect our teacher observations to the available frames of reference for judging teaching, we discover wide degrees of variation. The thing that stands out when we try to evaluate teachers is our inability to formulate identity statements. Our evaluative activities usually end with more or less precise statements about what we can say about the relationship between what we observe, our choice of a frame of reference, and our purpose.

What seems to be the case is that we are still searching for a "good" evaluation method—one that will allow us to increase our understanding of or reduce our uncertainty about—the evaluation of teachers before we have to make our evaluative statements. The search for a good evaluation then centers on an appropriate process for arriving at judgments, given that the judgments are not trivial in the first place. The importance of the frame of reference we choose is determined by our experience and then by showing its connectedness to our purpose before making a judgment. It should also be pointed out at this time that if we want our evaluative statements to be accepted by others, our evaluative purposes and activities have to be accepted by that "public." Evaluative purposes as well as evaluative activities that are not shared are simply passed over in silence.

Reliability/Validity

Much effort has been devoted to the development of particular measures for the evaluation of teachers. These measures allow us to feel comfortable about the variability in a given score or set of scores by providing reliable measures. When presented with reliable scores, we may assume a certain sense of "truth value" that is based upon their stability.

Stable scores that connect our observations to our frames of reference are said to be *valid*. Mary's evaluations as a teacher are valid to the extent that they are expressions of (measures shown to be related to) the publicly accepted definition (frame of reference) for the evaluation of pupils. It is interesting to point out again that reliability and validity are connected in measurement science in such a way that measures of reliability limit measures of validity. The stability of our observations limits their usefulness in providing meaning.

Valid scores are important for the simple reason that they have value in exchange. When there is agreement that the scores are valid (have referent value), they can be used as substitutes (proxies) for comparisons with the frame of reference chosen. Mary's evaluations of her pupils were significant somethings to the extent that they were both reliable and valid. They were important to the extent that they could serve as proxies for reading ability in the evaluation of her pupils. As an aside, it is pointed out that our present teacher evaluation methods often face the same concerns expressed by pupils concerning their evaluations. These concerns focus on our inability to establish any sense of confidence in the relationship posited between the measures and the purpose of the effort.

Three other concerns that also limit the usefulness of teacher evaluations not covered under the concepts of reliability and validity are: (1) the act of observation itself perturbs (distorts) our observations (the Heisenberg principle); (2) the observations perturb the observer; and (3) the noninvolved but influenced audience (those who judge the value of the statements) may not share our purpose of the activity.

Making sense out of data, that is, the transformation of data into information in order for our evaluative statements to have a shared meaning is therefore, at best, a precarious venture. Evaluative activities seem to more closely resemble "taking a walk" than they do "following a road."

Part II

Next, we will present some do's and don'ts for structuring teacher evaluations designed to improve instruction. A general model of evaluation that is intended to be more consistent with the content and purpose of instructional improvement is then presented. An evaluative strategy follows and conclusions about the evaluation of teaching as a professional activity are presented.

What Not to Do

In the evaluation of teaching, we often hear that we should focus on student learning. Learning (knowledge acquired through systematic study) is of interest to most people and they also tend to feel that knowledge is not trivial. It is worth evaluating. But, in order to use pupil learning in evaluating teachers, the relationship between teaching and learning has to be assumed to be at least *reversible* or *deterministic* or *necessary* or *sequential*. By assuming any or all of these forms of interrelatedness, it follows that teaching can be evaluated on the basis of student learning. How often have you heard the statement: "If the student hasn't learned—the teacher hasn't taught" to express the relationship between teaching and learning?

We know that the relationship between teaching and learning is not necessarily *reversible*. This says that because "I taught it *and* you know it" means the same thing as "since you now know—I taught it to you." Hard to believe. The *deterministic* assumption is also questionable. Who, as a teacher, would be willing to state that *"only* because I taught it—you know it." The *necessary* and *sequential* assumptions are also questionable. We know that for much of what is learned, a teacher is not needed and we all would agree that learning does not necessarily follow from teaching. What some pupils learn, they learn "in spite of the teacher." The problems associated with these assumptions are clarified by shifting the focus to the evaluation of a particular teacher or set of teachers. One might ask: "Is it fair to say that *a* teacher is responsible for (taught) what someone knows?" I think not. Anyone who has been evaluated as a teacher on measures of student learning should quickly learn to encourage only the "bright" students to take the test. The point is that the relationship between pupil learning as the measure of a teacher's effectiveness as a teacher is questionable.

A second common mistake we often make when we evaluate teachers is to focus on measures of teacher "characteristics." At the superficial level, we obtain measures of dress, speech, and bearing. It is not difficult to point out that a well-dressed, articulate, erect teacher may not be teaching anything. The measures are trivial. So much for the superficial. The more sophisticated measures of teacher characteristics available are those with a psychological and/or a sociological base. The concern in their development has been with the search for the "teaching personality" and/or the "teaching climate."

The interrelatedness of these measures appears to hinge on assumptions that are either *contingent* or *stochastic* or *coextensive* or *sufficient* when establishing a relationship between "characteristics" and "good" teachers for our evaluations (Note 2). The relationship is assumed to be *contingent* in the sense a "right" personality exists for teaching; *stochastic* in that a "good" person is *probably* also a good teacher; *coextensive* in that a "good" person *is* also a good teacher; and *sufficient* in that a "good" person can be a good teacher *regardless of anything*.

After years of research, the search for the "teaching personality" remains only a search. We have found that good persons do make good teachers but so do average and less-than-average persons when available psychological and/or sociological measures are employed as criteria. Maybe we persist in our search for the teaching personality simply because it is easier to work with people who have "good" personalities and are "interesting" in their interpersonal relationships. The evidence to date points out that a good person does not necessarily make a good teacher!

Third, the sought-for relationship between teaching and learning becomes more defensible the further one examines the possible ways that interrelatedness may exist. It appears that the teaching/learning relationship is at best *irreversible* or *substitutable* or *interdependent*. Perhaps the relationship is most clearly expressed when we are willing to assume interdependence. There appears to be little doubt that people do learn from each other, and what they learn forms a basis for additional learning. At least in the field of education, interdependence seems to express the case. The only measures of learning that would be acceptable (defensible) when evaluating teachers would be those that express what was learned by both parties. To find an appropriate learning measure for evaluating teachers, we should focus our search on what we know about what one would be expected to know in order to "practice teaching" as a professional activity.

What to Do

We will begin to focus the search for a way to improve our evaluation of teachers for the purpose of improving instruction by elaborating what we know about instruction. Our "theory of instruction" (following Bruner, Note 3) will hopefully provide us with a more direct measure of teaching skills.

The Assumptions

Teacher evaluations that serve our purpose are, of necessity, based upon a triadic relationship between what is known about teaching, that is, a *theory of instruction*, what is known about this particular person as a teacher, and our evaluative statements. A teacher is said to be a person with a theory of instruction in hand. A theory of instruction is prescriptive. It is the "tool" that is necessary for knowing what to do (how to behave) as a professional teacher. It is not the same as a theory of how students learn.

Instructional theory deals with what we know about predisposition, structure, sequence, and the provision for consequences. Under predisposition, we are concerned with developing acceptable measures that answer questions about what teachers do to address their pupils' "readiness" for teaching. These teaching skills establish a favorable "climate" toward the use of the mind. Examples of such questions might include: Does the teacher provide examples, pretest, state prerequisite knowledge, minimize risk, provide for exchange in relationships, create skepticism?

Under structure, the concern is with the development of acceptable measures for answering questions of subject matter organization. Teachers are expected to provide students with knowledge that is economical, productive, and powerful. Examples might include: Does the teacher state the assumptions, simplify diversity in content, show how to increase productivity, evaluate assumptions, state the general propositions, test relationships?

Sequence is the teaching skill that focuses on the activities necessary for subject matter presentation. Sequencing is used in an effort to demonstrate knowledge of the rules associated with definition, representation, manipulation, decomposition, recomposition, and so forth, for a variety of symbol systems that are known (have high probability) to improve a learner's ability to understand and explore. Questions might include: Does the teacher define, address order, represent concepts, use media, contrast inductive-deductive reasoning, vary pace, review?

Consequences have to do with a teacher's use of reinforcement. Consequences, as teaching activities, complete the formal teaching cycle. The learner is provided with a predisposition to continue to be a learner. The evaluative concern is with the development of measures that would increase our ability to answer questions about teachers. As examples: How is success treated? What risks are allowed? Who control rewards? How is support provided? Are measures related to tasks? How frequent are measures taken?

Conclusions

This paper began by assuming that the evaluation of teachers is an administrative concern. In the beginning, it was stated that most administrators would agree that teaching is poorly evaluated. The case was later made that this feeling could be changed if our evaluative efforts were made clear. The process that was elaborated included a focus on an evaluation strategy to serve the purpose of improving instruction. Before we addressed this problem directly, we clarified our understanding of evaluation and evaluative activities. The theory of evaluation presented was tied to the evaluation of teachers by establishing a relationship between the improvement of instruction as a purpose and evaluative activities. Our understanding of both evaluation and the evaluation of teaching would be improved if we agreed on a theory base for our discussion. Our task in evaluating teachers is to end with statements that relate the behavior of teachers to an accepted theory of instruction. The focus of this theory, it was argued, should not be on measures of student learning or on the personal/social characteristics or affinities of teachers. It was left to the readers to decide how to conduct their teacher evaluation instruments and who should be involved in the evaluation activity.

It is hoped that this discussion has increased your tolerance for the ambiguity inherent in the evaluation of teachers and will lead you to think differently about how you make evaluative statements. The intent was to encourage you to dream of alternatives to present teacher evaluation concerns by challenging how you think about and what is meant by the concept "evaluation." When someone asks: "How do you evaluate a teacher?" we hope you will respond: "It all depends on the purpose."

Note 1: This statement paraphrases the rule laid down by Vilfredo Pareto in *The Mind and Society* (New York: Harcourt-Brace, 1935). He presented the idea that the small number of persons (his elite) at the end of a distribution based on any criteria is of central interest in the study of composition. Par. 2027 and 2031.

Note 2: My first exposure to the underlined terms in this section was in Hans Zetterberg's *On Theory and Verification in Sociology* (Stockholm: Almquist and Wiksell and New York: Tressler Press, 1954).

Note 3: Jerome S. Bruner, *Toward A Theory of Instruction* (Cambridge, Mass.: Belknap Press of Harvard University, 1966).

Dr. Foley is Professor and Chairman of the Division of Educational Administration at The University of Iowa, Iowa City, Iowa. He has also served as Director of the Iowa Educational Information Center and the Student Development Center at The University of Iowa. He currently mediates for the Iowa Public Relations Board. Professor Foley has conducted research in the areas of educational decision-making, evaluation, theory, and is actively involved in leadership training.

THE EXECUTIVE REVIEW
Published eight times a year by

The University of Iowa
210 Lindquist Center
Iowa City, Iowa 52242

Institute for School Executives

William C. Bozeman, *Editor*
George A. Chambers, *Director*